*Proceedings 2nd World Congress of African Linguistics, Leipzig 1997*
H. E. Wolff and O. Gensler (eds), pp. 33-47, Köln: Rüdiger Köppe. 2000

# WRITING A REFERENCE GRAMMAR OF AN AFRICAN LANGUAGE: CONCEPTUAL AND METHODOLOGICAL ISSUES*

PAUL NEWMAN

One of the important tasks of African linguists is to document the languages of the continent through the preparation of dictionaries, text collections, and grammars. Many of the grammars constitute the first description of a previously unknown language, being based on original field research. (These are often Ph.D. dissertations written by younger scholars.) Others are general reference works on languages such as Fulani, Yoruba, or Zulu which are relatively well-known and which can boast an extensive literature. Regardless of the nature of the language concerned, the descriptions typically carry similar titles such as *A Grammar of the X Language*.

Having spent the past half-a-dozen years working on a comprehensive reference grammar of Hausa (Newman 2000), I have belatedly come to the realization that a reference grammar is not simply a descriptive grammar of a larger scope. Rather, it is quite different in kind from a "field grammar", and these differences relate to the nature of the methodology involved as well as the overriding principles governing the conceptualization and presentation of the final publication.

## 1. REFERENCE GRAMMARS

What is it that defines a reference grammar as compared with an initial descriptive grammar? Generally a reference grammar is a summary of knowledge about a large language on which work has been done by a large number of scholars over a considerable period of time. If done well, these grammars tend to be long-lasting and authoritative. For example, a century later, scholars still consult major 19th-century classics such as the Akan grammar written by Christaller (1827-1895) (Christaller 1875) or the grammar of Arabic written by Wright (1829-1898) (Wright 1955). Thus, in preparing a reference grammar, one has an enormous responsibility to achieve high levels of completeness and accuracy. One also has practical matters to keep in mind.

### 1.1. Be timeless (i.e. write for posterity)

Well-constructed pedagogical courses for teaching Hausa or Swahili or Bambara to non-native speakers aim to achieve naturalness. The goal is functional proficiency, which involves being pragmatically appropriate and culturally up to date. Good, lively materials, whether graded lessons, introductory readers, or practical dictionaries, need to have natural, pertinent examples. As noted above, reference grammars, on the other hand, tend to last a long time. One should thus avoid reference to topical matters such as the fall of the Berlin wall, Boycott's century accomplished in the test match against the West Indies, or Nelson Mandela's release from prison and the collapse of the apartheid system. One doesn't want a book that is antiseptic and dull; but one doesn't want one that is so loaded with the events of the moment that it soon appears out of fashion and awkwardly dated.

Similarly, one should avoid current theoretical debates as much as possible. Ten years from now, the reader won't have the slightest idea what so perturbed you about the obligatory contour principle, why you went on at such length about this type of rule ordering versus that, or why you adopted such a polemic tone in arguing for a particular principle of feature geometry as opposed to some other (probably propounded in a now forgotten MIT Ph.D. dissertation). One of course has to make use of theoretical ideas and metalinguistic concepts that fundamentally help with one's analysis – a truly atheoretical work consisting of a mishmash of unconnected facts would serve no purpose – but the theory needs to be digested and incorporated implicitly in the natural flow of the description and not be overly highlighted with neon lights or presented in an argumentative and bombastic spirit.

## 2. WRITE FIRST

The key to preparing a reference grammar in a sound and efficient manner is to write *before* you do your research. In embarking on a reference grammar project, you of course need to do background reading in order to get a sense of what has been done on the language. You also want to familiarize yourself with previously published grammars of different types and on different languages. But as far as your specific reference grammar is concerned, writing should precede research.

I can imagine that someone will immediately raise the objection that it is necessary to do the research first; otherwise, how can you know what to write about? This objection may be valid for work on an unknown language, but here is where a reference grammar is different: Anyone who proposes to do a reference grammar on a major language should know enough about the language to sit down and write a first draft of the book without notes. If the person cannot do this, he or she shouldn't be tackling the grammar!

If one concedes that a competent scholar *could* start writing at the outset with the research to follow, there is still the question of why one *should* adopt this reverse order. There are actually a considerable number of reasons, but I will limit myself to mentioning two, concerning presentation and discovery.

### 2.1. Presentation

First, the reverse order highlights the fact that in preparing a reference grammar, *major* problems throughout the work have to do with presentation. What goes where? How much detail? Should one include historical information? How about dialectal information? How much can one assume that the reader knows? The problems appear to be never-ending, but they cannot be avoided. Scholarly articles with new data or important insights can achieve their objectives even if the presentation is not ideal. But a reference grammar that is not "user-friendly" risks sitting idly on the shelf and not being used. What this means is that one needs to keep one's ultimate audience in view throughout the project. This is not something that can wait until the late stages of formatting the final version.

#### 2.1.1. The audience

One problem relates to the various constituencies that have to be taken into consideration. As far as scholars are concerned, in preparing a reference grammar one has to come to a compromise between the conventions of one's particular area of specialization and the conventions and knowledge of the general linguist. For example, when I undertake to write a technical article on Hausa for *Studies in African Linguistics*, *Journal of African Languages and Linguistics*, or *Afrika und Übersee*, I am able to address Hausaists who I can assume already have an understanding of the basics of Hausa. A reference grammar of Hausa, on the other hand, is different. The intention is that it will be used not only by Hausaists and Chadicists, and not even only by Africanists, but also by general linguists who may want access to information on Hausa for typological, comparative, or theoretical studies. These general linguists need guidance and explicit explanation, whether having to do with tone marking or the availability of morph-by-morph glosses.

An equally (if not more) important matter is keeping the native-speaker community in mind. We European and American linguists sometimes forget about the realities of the modern world and slip into the anachronistic idea that the languages we are studying are "exotic" languages spoken by a handful of people somewhere off in the bush. But this couldn't be further from the truth. In the case of Hausa (as with Swahili, etc.) we are dealing with a major world language, probably spoken by more people than speak European languages such as Czech or Danish. It is a language that is used in the schools and the mass media, one with established intellectual and cultural traditions, and one that has many native-speaker scholars with Ph.D.s from major universities around the globe. Given these realities, it is a serious error to work in a Eurocentric vacuum.

The clearest problems concern orthography, where the academic tradition has been to use linguistic transcription for non-Western (African) languages even when the languages have a well-established orthography of long standing. (By contrast, linguistic works on European languages generally use the established writing system of the language in question.) Admittedly, a linguistic transcription that consistently notes the phonemic elements and contrasts in a language is scientifically preferable and more informative for the linguist than a work using popular orthography; but one cannot totally ignore the fact that native speakers will want to consult the grammar and thus will not want to see their language written in a way that they cannot recognize. In certain instances, for example with regard to tone or accent which is not reflected in the orthography, the linguist may need to deviate from the conventional writing system. But in general, there has to be some kind of compromise between what one ideally would like as a linguist and the accepted orthography of the language being described.

In Hausa, for example, I have thus gone along with established orthographic convention in leaving word-initial glottal stop unwritten (e.g. **askā** ('razor') not **'askā**), even though /'/ is phonemic in the language and this orthographic convention masks the important generalization that all words in Hausa begin with a consonant. Similarly, I have employed the orthographic digraph **ts** rather than /s'/ for the ejective sibilant, and the orthographic **'y** rather than /ɗy/ for the palatalized /ɗ/. The diphthongs **ai** and **au** pose no problem for me, since I view them as complex nuclei; but I would contend that even people who phonemically analyze them as /ay/ and /aw/ should follow the orthographic rules and transcribe them as **ai** and **au**.

In the case of vowel length, which plays such an important role in Hausa (as does tone), I felt that it was essential for it to be marked in the grammar even though it is ignored in the established Roman orthography. (It is indicated in **àjàmi**, the Arabic-based script.) However, in order that my transcription be less jarring and offensive, I chose to mark vowel length not by means of double vowels (e.g. **aa**, **ee, uu**), as is typically done in linguistic articles on the language, but rather by means of a macron (e.g. **ā, ē, ū**), a diacritic that could be ignored by the native speaker, e.g.

| | | | | | |
|---|---|---|---|---|---|
| (1) | **kujḕrā** | (orthographic **kujera**) | *not* | **kujèeraa** | 'chair' |
| | **abūbuwǎ** | (orthographic **abubuwa**) | *not* | **'abuubuwàa** | 'things' |
| | **Hàusǎwā** | (orthographic **Hausawa**) | *not* | **hàwsàawaa** | 'Hausas' |
| | **tsaikǒ** | (orthographic **tsaiko**) | *not* | **s'aykòo** | 'roof frame' |
| | **'yātā** | (orthographic **'yata**) | *not* | **ɗyaataa** | 'my daughter' |

Concerning word division, I also generally went along with the orthographic rules rather than follow my intuition as a linguist. For example, the initial negative marker **bà** (with a short vowel), which is used in the

completive, and the future marker **zā** cliticize to and form a disyllabic word with the following weak subject pronoun, e.g. **bàtà dāwō ba** 'She didn't return'; **bà'à hanà̄ mu ba** 'One didn't prevent us'; **zāsù tàimàki Mūsā** 'They will help Musa'. However, in keeping with accepted practice, I transcribed them throughout the grammar as separate words, i.e. **bà tà**, **bà à**, and **zā sù**, respectively, rather than impose my scholarly viewpoint on the language.

### 2.1.2. The format

Since reference grammars tend to be large works, one faces special organizational problems concerning the order of presentation. What parts of the grammar must readers already have read before they can understand later chapters and sections? How can they locate specific material that they are interested in? After much thought, I adopted a somewhat novel format for my grammar in order to try to best address the needs of the reader. Let me describe what I did.

First, I took the notion of *reference* grammar seriously, i.e. apart from a handful of Hausa specialists who are devoted to the study of the language, the reality is that most people are not going to read it. The fact is that my *Hausa Grammar* is just not as exciting as a novel by John Grisham or Graham Greene! It's a depressing thought after putting so much effort into preparing a scholarly work to realize that it is not going to be read; but one has to take this as a given. What people will do is *consult* it, as one does in the case of an encyclopedia or a dictionary. Thus, unlike most grammars that begin at the beginning, e.g. phonology, and work up to more complex structures, I presented specific topics in independent chapters. As compared with Wolff's (1993) grammar, for example, which consists of seven chapters (plus the bibliography), my grammar contains 80 chapters, varying in length from 1 to 56 pages. Some of the chapters are quite general, e.g. equational sentences, gender, phonology, plurals, verbal nouns, whereas others are fairly specific, e.g. augmentatives, cognate accusative, expressives of contempt, proper names, remnant verb suffixes. Moreover, in a striking departure from the sequencing that one typically finds in grammars, the chapters are presented in alphabetical order; thus Chapter 1 is "Abstract nouns" and Chapter 80 is "Writing systems and orthography".

The other innovation in presentation has to do with notes. In close to 800 pages, there are no footnotes or endnotes. (To me nothing is more annoying than finding a footnote number in the middle of a sentence, whereupon one has to stick one's thumb in the page so as not to get lost while one looks for the notes some 37 pages later at the end of a chapter, only to find a note such as *ibid.* or *op. cit.* or, even worse, "This absolutely essential matter is explained in my unpublished M.A. thesis presented at Mt. Everest Community College (Nepal)".) Instead, borrowing from other works that I had seen – and here I followed the principle that one should always feel free to lift useful ideas from other works – I used notes *within* the

text, but set apart by special markers, indentation, and a smaller type size. Specifically, I limited myself to three different types of notes, analytical (AN), historical (HN), and dialectal (DN), indicated as in (2):

(2) °AN: Various scholars (*citations provided*) have suggested that **wà** is a bound inflectional suffix ...

◊HN: The **wà** variant is probably derived historically from the preposition **gà** (*citations provided*), the lenition to **/w/** coming about as a result of the move from the strong position following the direct object to the weak position after the verb ...

ΔDN: In the Bauchi dialect, spoken in the far southeastern edge of Hausaland (*citations provided*), noun indirect objects (marked by **wà** or **mà** + NP) occur after the direct object ...

### 2.2. Writing as discovery

Second, it is in the course of writing that one discovers what one knows and what one doesn't, i.e., what one has to investigate further. For example, on the basis of examples that all Hausaists are familiar with, grammars of Hausa normally describe a rule that changes **f** into **h** before back rounded vowels, e.g.

(3) **f → h** / ___V[+round]

| | | |
|---|---|---|
| (a) | **tàfi** | 'go', cf. **tahO** 'come' (with the **-O** 'ventive' ending) |
| | **dafà** | 'cook', cf. **dàhu** 'be well cooked' (with the **-u** "sustentative" ending) |
| | **tsòfàffī** | 'old' (pl.), cf. **tsōhō** 'old' (sg.) |
| (b) | **hòtō** | 'photo' |
| | **hōdà** | 'powder' (Note: Hausa has no **/p/** as opposed to **/f/**.) |

It turns out that this rule is much less productive than generally thought. In fact it is limited to a small number of high-frequency forms such as **tàfi** 'go' vs. **tahō** 'come'; otherwise /f/ before /o/ and /u/ normally stays /f/, e.g.

| | | |
|---|---|---|
| (4) | **nùfā** | 'head towards', cf. **nufō** 'head this way' |
| | **tsāfà** | 'squeeze', cf. **tsàfu** 'be squeezed' |
| | **mahaifā** | 'womb, birthplace', cf. **màhàifū** 'birthplaces' |
| | **ƙafà** | 'foot, leg', cf. **ƙafàfū** = **ƙafāfuwà** 'feet, legs' |
| | **kafā** | 'small hole', cf. **kafōfī** 'holes' (also **kafōfin wātsà làbārī** 'the media') |

But how would anyone have known in advance to investigate this question, since, given our familiarity with and conversational use of the high-frequency forms, we didn't know that there was in fact a question? What led me to doubt the received knowledge was the conflict between the **f → h** rule that I had originally written in the chapter on phonology and the examples that I had typed up when I was writing the section on

plurals with final **-uu**, and where I discovered a considerable number of what I thought at first were "exceptions", namely **f**'s that stayed **f** (see examples above in (4)).

### 3. CLIMB ON THE SHOULDERS OF YOUR PREDECESSORS AND BORROW UNASHAMEDLY

In preparing a reference grammar one must read, read, and read. My published bibliography of Hausa and Chadic (Newman 1996) includes some 1,000 entries on Hausa. Apart from the items written in Russian, a language which I unfortunately do not know, I have read almost everything included in the bibliography. I depended heavily on a handful of sources, namely Abraham's (1959) grammar, the two volumes of Parsons' papers compiled by Graham Furniss (Parsons 1981), and Wolff's important grammar (Wolff 1993). But in preparing a reference grammar, all studies of the language are potentially valuable. One particularly should not neglect early sources. Some of them may appear to be naive and wrong in detail, but in many cases they provide a fresh view on things that today no one would see because we take a lot for granted. As an example, Hausa has an active palatalization rule before front vowels, affecting voiced and voiceless coronal stops and fricatives. This is generally described as **s**, **z**, **t**, **d** → **sh**, **j**, **c**, and **j**, respectively (both **z** and **d** going to /j/ (= [zh] in Northern dialects)), e.g.

(5) **tāsà** 'bowl', cf. pl. **tāsōshī** (with the **ōCī** suffix, where C is a copy of the preceding consonant)
**cìzā** 'bite', cf. pre-direct object form **cìji**
**wutā** 'fire', cf. pl. **wutàcē** (with the **àCē** suffix, where C is a copy of the preceding consonant)
**yàřda** 'agree', cf. pre-indirect object form **yařjè**

Any student in an introductory phonology course could (and would) write a single rule handling all four consonants as a natural class. Interestingly, this turns out to be inexact. If one looks at mid-19th-century works such as Schön's grammar and dictionary (Schön 1862, Schön 1876), one finds that the rule only applied to three of the four coronal consonants, namely **s**, **z**, and **t**, but generally not **d**, e.g. (examples as given by Schön 1876, but with underline added):

(6) **kaḏa** 'crocodile', pl. **kadoḏi**
**kaẓa** 'hen', pl. **kaẓhi**
**biṣa** 'animal', pl. **bisaṣhe**
**saṭa** 'theft', verb form **saċe** 'steal'

Surprisingly, it turns out that non-palatalizing **d** is still common in contemporary Hausa. With some words palatalization takes place, e.g. **gidā** 'house', pl. **gidàjē**; **gàdā** 'inherit', **magàjī** 'heir', etc.; but with many it does not, both in roots and in derived forms. Moreover, if the **d** is preceded by **n**, it turns out that the norm

is *not* to undergo palatalization. In other environments, palatalization or the lack thereof is lexically specific, e.g.

(7) **gìndī** (not *__gìnjī__) 'base'; **landè** (not *__lanjè__) (used in **jàtan landè** 'type of cloth'); **kundī** (not *__kunjī__) 'pad, thesis'

**bidōdī** 'thatching needles' ← **bidà** sg.; **marȉdī** 'snatcher' ← **rȉdā** 'snatch'; **jìdē** 'transport' (pre-pronoun object form) ← **jìdā**

How one handles this kind of variation/indeterminacy in a synchronic reference grammar is an open question; however, if one had based one's description strictly on current received knowledge, the non-palatalizing **d**'s would have been relegated to footnotes as minor exceptions, when in fact they are quite normal and not exceptions at all.

Drawing on other works also extends to questions of format and presentation. In deciding how to prepare one's grammar, it is essential to read widely to see how other people have organized their reference grammars, whether they be of another African language or of a language of Europe or Asia. One should feel free to model one's grammar on someone else's. It is not necessary to reinvent a new format. If you find a grammar that you like – Mosel and Hovdhaugen's *Samoan Reference Grammar* (1992) is one of my favorites – there is no reason why you couldn't adopt the template as is. It turns out that I didn't do this, but in principle there would have been nothing wrong with adopting someone else's grammar and simply replacing the facts by those of the language that you are working on. Although I eventually came up with my own approach to the problem of presentation, outlined above, I did take many other grammars into account before coming to my final decision.

#### 4. BE ESPECIALLY CRITICAL AND OPEN-MINDED WHEN IT COMES TO YOUR OWN WORK

Scholars planning to do a reference grammar are probably not novices. They have presumably worked on the language for years and written extensively about it. They have probably had to defend various analyses against other viewpoints with the result that emotionally they have a vested interest in what has been published. This being the case, it is essential to step back and say: "OK, I wrote X, Y, and Z, and at the time I had good reason for taking the positions that I did and I am pleased that such and such scholars accepted my analyses; but, now that I have to commit myself for purposes of a reference grammar, which means for the next half-century, do these analyses really hold up?"

Reduplication in Hausa provides a good illustration of where I had to alter an earlier, seemingly solid analysis. Hausa has numerous quadrisyllabic noun plurals and pluractional verbs manifesting partial reduplication, e.g.

| | | | |
|---|---|---|---|
| (8) | **bakunkunà** | 'bows' | ← **bàkā** sg. + **-unà** pl. |
| | **sūnànnakī** (← ***sūnàknakī**) | 'naming ceremony' | ← **sūnā** sg. + **-akī** pl. |
| | **kaɗànɗanī** | 'shea nut trees' | ← **kaɗanyà** sg. (base **//kaɗanì//**) + - **a-ī** pl. |
| | **shāwàřwařī** | 'pieces of advice' | ← **shāwařà** sg. + - **a-ī** pl. |
| | **maƙalƙàlē** | 'get lodged' (many or often) | ← **maƙàlē** |
| | **tafařfàsā** | 'boil on and on' | ← **tafàsā** (syllable final **s** → **ř**) |
| | **tararràtsē** | 'smash all over' | ← **tartsè** (base **//taratse//**) |

In previous works I analyzed such items in terms of two-syllable reduplication to the right, accompanied by dropping of the stem-final vowel, this latter being the norm with Hausa suffixation. This is illustrated in (9), where the presumed copied part is underlined. (Tone is to be ignored for the moment.)

| | | | |
|---|---|---|---|
| (9) | **bakunkunà** | 'bows' | ← **bakun(à)-kunà** |
| | **shāwàřwařī** | 'pieces of advice' | ← **shāwàř(i)-wařī** |
| | **maƙalƙàlē** | 'get lodged (many or often)' | ← **maƙal(e)-ƙàlē** |
| | **tafařfàsā** | 'boil on and on' | ← **tafas(a)-fàsā** |

If one looks deeper into the language, however, one finds that this analysis leaves something to be desired. First, the tones are problematic. But more serious (at least with regard to the noun plurals) is that this suffixal analysis treats these reduplicated items as totally unrelated to the formation of other quadrisyllabic plurals, some of which are reduplicated and some not, e.g.

| | | | |
|---|---|---|---|
| (10) | **fikà̲fikī** | 'pair of wings' | ← **fiffikè** sg. |
| | **gùmà̲gùmai** | 'logs' | ← **gungumē** sg. |
| | **kwàřā̲kwàřai** | 'spindles' | ← **kwařkwařō** sg. |
| | **malè̲malī** | 'mounds of *tuwo* (basic food)' | ← **malmalā** sg. |
| | **tařè̲wadī** | 'mudfishes' | ← **tařwadā** sg. |
| | **garè̲kanī** | 'fenced-in gardens' | ← **garkā** sg. |

In these examples, one can see clearly that the pluralization rule involves the insertion of a long vowel in antepenultimate position. Some of the plurals also involve suffixation, but the long vowel (either **a̲** or **è̲**) is clearly infixed before the final two syllables. Now, if one compares plural words such as **kaɗànɗanī** or **shawàřwařī** with the tonally identical forms **tařè̲wadī** and **garè̲kanī**, it suddenly becomes clear that essentially the same insertion process is involved, namely insertion of a heavy syllable or syllabic nucleus in antepenultimate position. In the one case a long vowel is inserted, in the other a reduplicated CVC is inserted, but from a metrical / prosodic point of view the process is the same. Assuming that this is the correct analysis for words such as **shawàřwařī,** there is no reason why it shouldn't also hold for all the other plurals and pluractionals in (9); i.e., we can do away with two-syllable suffixal reduplication as a process and generate all of the examples with CVC infixal reduplication, e.g.

(11) [alternative analysis: infixal reduplication], cf. (8) and (9)

| | | |
|---|---|---|
| **kaɗ̲ànɗanī** | **shāwà̲ř̲wařī** | |
| **ba̲kunkunà** | **sūnànnakī** | (← ***sūà̲knakī** ) |
| **ma̲ƙal̲ƙàlè** | **ta̲fa̲ř̲fàsā** | (← ***ta̲fa̲sfàsā** ) |
| **tara̲rràtsē** | | (← ***tara̲tsràtsē**) |

**5. BE FLEXIBLE (ALLOW JUDICIOUS INCONSISTENCY IN ANALYSES)**

Generally speaking, we like our analyses of a particular language to be internally consistent. Sometimes, however, the only way to provide an adequate description of the facts of a language is to allow different analyses to coexist in different parts of the grammar. For example, the Hausa diphthongs [ai] and [au] have been subject to two alternative interpretations. Some scholars have analyzed them as consisting of the vowel /a/ followed by a coda consonant /y/ or /w/, i.e. /ay/ and /aw/. The position that I have always taken (see Newman and Salim 1981), and which I have adhered to in my *Grammar*, is that the diphthongs are long nuclei comparable to long vowels, i.e., they have the structure VV rather than VC. There many reasons for this, having to do with the fact that long vowels in Hausa typically behave as a class whether the vowels are monophthongs or diphthongs. Two examples will suffice here. First, in plural formation involving disyllabic words with high-high tone, those words with a diphthong in the first syllable pattern with those having a long vowel in adding a suffix **à̲yē** as opposed to those with VC in the first syllable, which use an internal **à̲-ē** formative, e.g.

(12) sg. / pl.:**bāmī / bāmà̲yē** 'novice'; **dūlū / dūlà̲yē** 'type of basket'
**ɓaurē / ɓaurà̲yē** 'fig tree'; **maisō / maisà̲yē** 'farm which has been left unworked'
cf. **maskō / masà̲kē** 'mallet'; **kurfō / kurà̲fē** 'whip, lash'

Second, if a consonant, such as the definite article (which in the masculine and plural is **'-n**, i.e. **-n** with low tone) or the genitive linker **-n** (which is toneless **-n**), is added to a syllable containing a long vowel, the vowel is automatically shortened, since Hausa does not allow long vowels in closed syllables. This rule CVVC → CVC operates whether the two V's are identical (i.e. monophthongal) or not (i.e. diphthongal). By contrast, a consonant cannot be added to a syllable that is already closed, since Hausa does not allow consonant clusters. To do this, one must either insert an epenthetic vowel or make use of a transition particle. Again we find that the diphthongs pattern with the long vowels (written here with double letters), e.g.

(13) CVVC → CVC
**sâa** 'bull', **sân** 'the bull', **duuluu** 'type of basket', **duulunsà** 'his basket'
**mâi** 'oil', **mân** 'the oil'; **kìbau** 'arrows', **kìbansà** 'his arrows'
cf. **bâm** 'bomb', **bâm ɗ̂in** 'the bomb' (with particle **ɗī**); **tēbùř** 'table', **tēbùřinsà** 'his table'

The above seems clear and convincing. However, in looking at ideophones, something different happens. A very common canonical pattern for ideophones is CVCVC, where the two short vowels are identical (tone is variable). The canonical shape CVCVV does not occur (apart from CVCVV-CVCVV reduplicated forms) except in the case of words with final diphthongs, e.g.

| | | | |
|---|---|---|---|
| (14) | (a) | **tiƙis** | 'intensity of tiredness' |
| | | **gùnùs** | 'describes a terrible stench' |
| | | **shìrìm** | 'shady (of shrub or tree)' |
| | | **sumul** | 'very well (of health), washed clean' |
| | | **rirìs** | 'indicates intensity of crying' |
| | | **kwatsàm** | 'suddenly' |
| | | **jùgum** | 'being sad, dejected, despondent' |
| | (b) | **fatau** (i.e. **/fataw/**) | 'emphasizes greenness (completely green or deep green)' |
| | | **gàlau** (i.e. **/gàlaw/**) | 'vacantly, with mouth wide open' |
| | | **sarai** (i.e. **/saray/**) | 'excellent' |

For descriptive purposes, I have grouped the diphthong-final words in (b) with the C-final words on canonical grounds, i.e., I have treated **fatau**, **gàlau**, **sarai** as **fataw**, **gàlaw**, **saray**, respectively, *contrary* to my own analysis. But what is the alternative? One can't simply ignore the ideophone examples and hope that they will go away. On the other hand, one can't really change the general phonological analysis of diphthongs, since this is supported so strongly throughout the language. The only reasonable solution is to allow the two conflicting analyses to coexist.

### 6. DON'T BE TOO ORIGINAL AND CLEVER

One's whole academic career is spent trying to come up with new and innovative ideas. Our objective is to be more clever than the next guy. In a reference grammar, on the other hand, one wants to work within the canons of a field as much as possible. For example, one of the essential organizing principles of the Hausa verb system for the past thirty-five years or so has been the organization into morphological categories termed grades (Parsons 1960). As I pointed out many years ago (Newman 1973), there are in fact problems with the grade system. However, it is so central to Hausa scholarship that I decided that it was essential in a reference grammar to take this system as a given, in the same way that someone writing a reference grammar of Zulu has to operate with noun classes. I did add a substantial section providing a historical critique of the grade system, but in presenting the verbs themselves, I took the grade system as the starting point. (At 56 pages, the chapter entitled "Verb grades" is actually the longest in the book.)

Having said that one should make an effort to avoid unfamiliar originality, I would contend that one should go ahead and break with tradition if necessary. That is, there are some places in the grammar where you will conclude that the previous analyses are inadequate and beyond repair, so to speak. In these cases,

you have to assert yourself as an authority and present the analysis that is in keeping with the facts, even if it runs against received knowledge. Let me give one example with regard to the Hausa infinitive.

The long-held view about Hausa, which is commonly repeated and about which I have seen no objections, is that Hausa does not have an infinitive. How this erroneous idea got so well established I cannot say, although I suspect it is related to the fact that Hausa doesn't have anything comparable to the English preposition 'to', which we associate with the infinitive. On the other hand it is clear, as pointed out in the standard reference grammar of English (Quirk et al. 1985), that English also has bare infinitives without 'to', as seen in sentences such as "What she likes to do is *cook* food" or "He had better *tell* them."

What do I mean by an infinitive? It is a VERB form (i.e. not a nominalized form such as a gerundive) which is not marked for tense or person / number. In the case of Hausa, this generally means a verb form not preceded by the "Person-Aspect-Complex" which includes the weak subject pronoun plus the TAM (e.g. **sukàn** 'they.habitual'). Inflected finite verbs are illustrated in (15) (the verb is underlined):

| | | |
|---|---|---|
| (15) | **zā tà gayà manà** | 'She will tell us' (lit. future she tell to us) |
| | **sun sācè shi** | 'They abducted him' |
| | **mukàn būdê ƙōfàř à ƙarfè takwàs** | 'We (habitually) open the door at 8 o'clock' |

In certain environments, which we will call non-finite environments, one either gets a verbal noun (examples in (16a)), i.e. an item with real nominal properties such as gender, or something that has the form of a finite verb (examples in (16b)), the choice depending on the verb class, e.g.

(16) (a) **sun ƙi jìfà-nsà** 'They refused to throw (lit. throwing) at him' (Note genitive pronoun **sà**, lit. 'his throwing'.)
**řùbùtu-n wàsīƙà yanà dà sauƙī** 'Writing letters is easy'
**bà mù san lōkàcin isôwa-řsà ba** 'We don't know his time of arriving' (with verbal noun **isôwā**) (cf. **bà mù san lōkàcîn dà zâi isō ba** 'We don't know the time when he will arrive' (with finite verb **isō**))

(b) **sun ƙi sācè shi** 'They refused to abduct him' (Note object pronoun **shi** 'him')
**tā iyà gayà manà** 'She is able to tell us'
**yā hau katangā don jēfà duwàtsū** 'He climbed the wall to throw stones'
**būdê ƙōfàř yanà dà wùyā** 'To open the door is difficult'
**ɓōyè kāyan màsàřūřī sukà yi** 'It was hoarding basic foodstuffs that they did'

The status of the underlined words in (a) is not problematic. All scholars agree that one has a verbal noun, i.e. a nominalized verb form (in these examples plus a genitive linker). The question is, what is the status of the underlined form in the examples in (b)? The analysis of most Hausaists has been that these are verbal nouns that *just happen* to be identical to the finite verb form. But if they look just like verbs, have final vowel length alternations just like verbs, take direct object pronouns rather than genitive pronouns just like

verbs, and behave in other ways just like verbs – they can occur before indirect objects, which verbal nouns cannot – it seems to me that the simple-minded analysis is to recognize that they are verbs. If it looks like a rose and smells like a rose, then your best bet is that it is a rose.

But now we come back to the definition of an infinitive, namely a non-inflected verb form. In a sentence such as **bū̀dê ƙōfàř yanà dà wùyā** 'To open the door is difficult', **bū̀dê ƙōfàř** is thus an infinitive phrase. Apart from the word 'to', which Hausa does not have, the parallelism with the English infinitive is quite striking.

But there is something else that one needs to say about Hausa infinitives – and here we will see that the scholars who were looking at these forms as verbal nouns were not so far wrong. Whereas **bū̀dê** 'open' is a straightforward verb without nominal features, the phrase as a whole is nominal, i.e. **bū̀dê ƙōfàř** is syntactically a masculine NP. Consider the examples in (17), repeated from (16b), where the bracketing is essential.

| | | |
|---|---|---|
| (17) | [[**bū̀dê**]$_V$ [**ƙōfàř**]$_N$ ]$_N$ **yanà** … | 'To open the door …' |
| | **sun ƙi** [[ **sācè**]$_V$ [**shi** ]$_N$ ]$_N$ | 'They refused to abduct him' |
| | [[**ɓōyè**]$_V$ [**kāyan màsàřūfī** ]$_N$ ]$_N$ **sukà** … | 'To hoard basic foodstuffs …' |

To generalize, then, what we can say is that Hausa *does* have infinitive phrases and they have the structure given in (18):

(18) INFINITIVE = [V + obj]$_N$ (where V is a verb (*not* a verbal noun) and obj is d.o. or i.o.)

Two things about this structure are essential. First, note, as explained above, that the V in the infinitive phrase is a verb, *not* a verbal noun. Second, note that in Hausa – and this is where Hausa is different from English and many European languages – one *must* have an object in this construction. If the object is not present, one gets the following structure, which happens to be the structural definition of a verbal noun:

(19) [V]$_N$

It follows automatically from the object requirement that intransitive verbs do not have infinitives, e.g.

(20) **sun fārà** [[**zaunàwA**]$_V$]$_N$ 'They began sitting' (verbal noun), cf. **sun** [**zaunà**]$_V$ 'They sat' (finite verb)

Now note what happens to sentences with infinitive phrases if the object is deleted. In this case, the bare verb becomes dominated by N and is thereby automatically altered into a corresponding verbal noun,

usually with the weak verbal noun suffix **'-wā**. (Note, interestingly, that verbal nouns with **'-wā** are feminine whereas the source infinitive phrases are masculine.) Examples:

(21) [[**būdề**]$_V$ [**kōfàř**]$_N$ ]$_N$ → [[**būdề**]$_V$ [Ø]$_N$ ]$_N$ → [[**būdề**]$_V$]$_N$ → **būdềwā** 'opening',
e.g. **būdềwā tanà dà wùyā** 'Opening (it) is difficult'
[[**jēfà**]$_V$ [**duwàtsū**]$_N$ ]$_N$ → [[**jēfà**]$_V$ [Ø]$_N$ ]$_N$ → [[**jēfà**]$_V$]$_N$ → **jēfàwā** 'throwing'
e.g. **yā hau katangā don jēfàwā** 'He climbed the wall in order to throw (them)' (lit. for throwing)
[[**gayà**]$_V$ [**manà**]$_N$ ]$_N$ → [[**gayà**]$_V$ [Ø]$_N$ ]$_N$ → [[**gayà**]$_V$]$_N$ → **gayàwā** 'telling'
e.g. **tā iyà gayàwā** 'She is able to tell (us)'

Scholars have long wondered why it is that in non-finite environments, one gets a verb without **'-wā** when there is a following object, but a form with **'-wā** when there is no object. The answer is that the form without **'-wā** is a simple verb used in an infinitive phrase, whereas, since Hausa does not allow bare infinitives, the verb form without an object is obligatorily altered into a corresponding verbal noun.

## 7. CONCLUSION

A good reference grammar must be accurate, it must be comprehensive, it must be coherent, and it must be well written and accessible to the reader. The most important requirement, however, is that it must be finished! Reference grammars invariably turn out to be much more difficult and much more time-consuming projects than anyone could ever anticipate in advance. The result, as I learned from painful first-hand experience, is that the preparation of a serious reference work tends to drag on and on, the end-point constantly changing its position like a mirage. In the book *Caught in the Web of Words* (Murray 1977), there is a wonderful photograph of James A.H. Murray, the great English lexicographer, with a long white beard surrounded by his coworkers. When Murray died in 1915, he had been working on the OED (Oxford English Dictionary) for 36 years, and yet it was still far from completion. Because of the nature of the OED project, it was eventually finished some years later, but most of us cannot count on being that lucky – our skills and experiences are much too specialized, and an individual scholar cannot readily be replaced. We thus need to force ourselves to bring our work to culmination and avoid the temptation to check the transcription one more time, to insert a cogent example, or to add a discussion to take into account a new article that has just appeared. At some point one simply has to say, "For better or worse, this is it." Finishing a reference grammar and getting it published is essential in order to disseminate our knowledge to others as well as to have the pleasure of seeing the product of our extensive labors appear while we are still alive and well.

## REFERENCES

Abraham, R.C. 1959. *The Language of the Hausa People*. London: University of London Press.

Christaller, Johann G. 1875. *A Grammar of the Asante and Fante Language Called Tshi*. Basel: Basel Evangelical Missionary Society.

Mosel, Ulrike & Even Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Murray, Katharine M.E. 1977. *Caught in the Web of Words: James A. H. Murray and the Oxford English Dictionary*. New Haven: Yale University Press.

Newman, Paul. 1973. Grades, vowel-tone classes and extensions in the Hausa verbal system. *Studies in African Linguistics* 4:297-346.

———. 1996. *Hausa and the Chadic Language Family: A Bibliography*. (African Linguistic Bibliographies, 6.) Cologne: Rüdiger Köppe.

———. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.

Newman, Paul & Bello Ahmad Salim. 1981. Hausa diphthongs. *Lingua* 55: 101-21.

Parsons, F.W. 1960. The verbal system in Hausa. *Afrika und Übersee* 44: 1-36.

———. 1981. *Writings on Hausa Grammar: The Collected Papers of F. W. Parsons*, edited by Graham Furniss. Ann Arbor, MI: UMI Books on Demand.

Quirk, Randolph et al. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

Schön, James Frederick. 1862. *Grammar of the Hausa Language*. London: Church Missionary House.

———. 1876. *Dictionary of the Hausa Language, with Appendices of Hausa Literature*. London: Church Missionary House.

Wolff, H. Ekkehard. 1993. *Referenzgrammatik des Hausa*. (Hamburger Beiträge zur Afrikanistik, 2.) Münster & Hamburg: LIT.

Wright, William. 1955. *A Grammar of the Arabic Language*. 3rd revised edition. Cambridge: Cambridge University Press.

## ACKNOWLEDGMENTS

* Preparation of the Hausa Reference Grammar described in this paper was supported by grants from the U.S. Department of Education (P0-17A10037), the National Endowment for the Humanities (RT-21236), and the National Science Foundation (DBS-9107103). Travel to the Congress was made possible through the generosity of the Indiana University Overseas Conference Fund of the Office of International Programs.